

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出 願 年 月 日
Date of Application:

2000年 7月26日

願 番 号
Application Number:

特願2000-231346

願 人
Applicant(s):

株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

USSN 09/808951

MATTINGLY, STANGER, MALUR
+ BRUNDIDGE

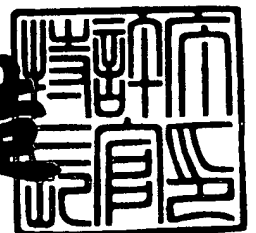
(703) 684-1120

DKT: NIT 266

2001年 3月 9日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 NT00P0425

【提出日】 平成12年 7月26日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 9/46

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内

 【氏名】 森 利明

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内

 【氏名】 山▲崎▼ 康雄

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100068504

 【弁理士】

 【氏名又は名称】 小川 勝男

 【電話番号】 03-3661-0071

【選任した代理人】

 【識別番号】 100086656

 【弁理士】

 【氏名又は名称】 田中 恭助

 【電話番号】 03-3661-0071

【選任した代理人】

 【識別番号】 100094352

 【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 仮想計算機システムにおける計算機資源の割当て方法および仮想計算機システム

【特許請求の範囲】

【請求項 1】 仮想計算機システムにおいて、現用系 VM と待機系 VM とを用意し、現用系 VM の OS にはあるアプリケーションプログラムを実行するに十分な主記憶領域を割り当て、待機系 VM の OS には当該アプリケーションプログラムの実行には不十分な小さい主記憶領域を割り当て、前記現用系 VM において障害が生じたとき前記現用系 VM に割り当てられた主記憶領域の一部または全部を待機系 VM に組み入れることを特徴とする仮想計算機システムにおける計算機資源の割り当て方法。

【請求項 2】 更に、前記あるアプリケーションプログラムと同様の業務を遂行するホットスタンバイ用アプリケーションプログラムを補助記憶装置に用意し、前記現用系 VM において障害が生じたとき前記あるプログラムに割り当てられた主記憶領域を待機系 VM に組み入れると共に、前記ホットスタンバイ用アプリケーションプログラムを待機系 VM の OS に割り当てられた主記憶を利用して実行することを特徴とする請求項 1 記載の仮想計算機システムにおける計算機資源の割り当て方法。

【請求項 3】 前記仮想計算機システムは複数の OS を制御する仮想計算機モニタを有し、主記憶の OS、仮想計算機モニタ、アプリケーションプログラム、及び未使用域ごとの使用量と、使用している OS または仮想計算機モニタと、及び障害時の切り替え先の OS または仮想計算機モニタを記録した資源管理テーブルを用意し、仮想計算機モニタは現用系 VM において障害が生じたとき前記資源管理テーブルを参照し、必要な主記憶領域について前記使用量だけ、記憶された障害時の切り替え先に主記憶領域の組み込みを指示する通知を出すことを特徴とする請求項 1 記載の仮想計算機システムにおける計算機資源の割り当て方法。

【請求項 4】 仮想計算機システムにおいて、現用系 OS と待機系 OS とを設け、前記現用系 OS に割り当てた主記憶の利用用途を記録し、前記現用系 OS 上で動作するアプリケーションプログラムに異常が発生した場合、前記記録した利用

用途を参照して異常が発生したアプリケーションプログラムが利用した主記憶を前記待機系OSに割り当て直すことにより前記待機系OS上で異常が発生したアプリケーションプログラムと同様の業務を遂行するアプリケーションプログラムを実行することを特徴とする仮想計算機システムにおける計算機資源の割り当て方法。

【請求項5】複数のOSを制御する仮想計算機モニタを有する仮想計算機システムにおいて、現用系OSは障害時に待機系OSの下で同様の業務を遂行するプログラムが実行されるホットスタンバイジョブのアプリケーションプログラムを含む1または複数のアプリケーションプログラムの実行に使用する計算機資源を計算し、資源が充足している場合は前記仮想計算機モニタにどのアプリケーションプログラムがどの資源を使用しているかを通知し、資源が不足している場合は前記仮想計算機モニタから新たな資源を確保し、現用系OSまたは前記ホットスタンバイジョブのアプリケーションプログラムに障害が発生したとき、前記仮想計算機モニタは少なくとも前記ホットスタンバイジョブのアプリケーションプログラムが使用していた資源を待機系OSに組み込む通知を待機系OSに出すことを特徴とする仮想計算機システムにおける計算機資源の割り当て方法。

【請求項6】複数のOSと複数のOSを制御する仮想計算機モニタと主記憶装置を含む資源を有する単一の計算機システムを用いて、同様の業務を遂行する複数のプログラムを排他的に実行するための方法であって、

前記複数のOSのうち、第1のOSにおいては、第1のOS上で動作する第1のアプリケーションプログラムに割当てた資源を前記仮想計算機モニタに通知し

第1のOSが第1のアプリケーションプログラムの障害を検知すると前記仮想計算機モニタに通知し、前記仮想計算機モニタにおいては、第1のOSからの障害の検知の通知を受けると第1のアプリケーションプログラムが使用していた資源を第1のOSから切り離せしめ、当該資源を第2のOSに割り当て、第2のOSに第1のアプリケーションプログラムと同様の業務を遂行する第2のアプリケーションプログラムの起動を通知し、前記第2のOSにおいては、第2のアプリケーションプログラムの起動によって使用される資源を前記割り当てられた資源か

ら割り当てることを特徴とする仮想計算機システムにおける計算機資源の割り当て方法。

【請求項 7】複数の OS と複数の OS を制御する仮想計算機モニタと主記憶装置を含む資源を有する単一の計算機システムを用いて、同様の業務を遂行する複数のプログラムを排他的に実行するための方法であって、

前記複数の OS のうち、第 1 の OS においては、第 1 の OS 上で動作する第 1 のアプリケーションプログラムに割り当てた資源を該仮想計算機モニタに通知し、前記仮想計算機モニタが第 1 の OS の障害を検知すると、第 1 の OS が使用していた資源の一部または全部を第 2 の OS に割り当て、第 2 の OS に第 1 のアプリケーションプログラムと同様の業務を遂行する第 2 のアプリケーションプログラムの起動を通知し、第 2 の OS においては、第 2 のアプリケーションプログラムの起動によって使用される資源を前記割り当てられた資源から割り当てることを特徴とする仮想計算機システムにおける計算機資源の割り当て方法。

【請求項 8】第 1 の OS と第 2 の OS がホットスタンバイ構成にある場合、第 1 の OS 上で動作するアプリケーションプログラムの障害を検知すると、障害のあった第 1 のアプリケーションプログラムがホットスタンバイジョブであったときのみ第 2 のアプリケーションプログラムを第 2 の OS 上で動作させることを特徴とする請求項 6 記載の仮想計算機システムにおける計算機資源の割り当て方法。

【請求項 9】仮想計算機システムにおいて、現用系 OS と待機系 OS と複数の OS を制御する仮想計算機モニタを備え、前記現用系 OS は少なくとも前記現用系 OS の下で実行されるアプリケーションプログラムの障害度を監視し障害が回復困難であるときその旨を前記仮想計算機モニタに通知する障害度通知処理部と、前記仮想計算機モニタから現用系 OS に割り当てられた資源の切り離し通知を受けると要求された資源を切り離す資源切り離し処理部とを有し、前記待機系 OS は前記仮想計算機モニタから新たに組み入れるべき資源の通知を受けて通知された資源を自 OS に組み入れる資源追加処理部を有し、前記仮想計算機モニタは現用系 OS の障害を検知する OS 障害検知処理部と、現用系 OS の障害を検知すると待機系 OS に移行すべき資源を求め待機系 OS に新たに組み入れるべき資源

を通知する手段と、現用系OSからアプリケーションプログラムの障害の通知を受けて切り離すべき資源を求めて現用系OSに通知する手段と、資源の切り離しの完了後待機系OSに新たに組み入れるべき資源を通知する手段とを有することを特徴とする仮想計算機システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、計算機システムに係り、特に、1台の計算機上で複数のOSが動作する計算機システムにおける各OSに対する計算機資源の分配方法に関する。

【0002】

【従来の技術】

1台の計算機システム上で複数のオペレーティングシステム(OS: Operating System)を実行させるシステムとしては、仮想計算機システム(VMS: Virtual Machine System)がある。仮想計算機システムは、計算機システムが保有するプロセッサ、主記憶装置、補助記憶装置、通信制御装置等の物理資源を論理的に分割して複数の論理的な仮想計算機(VM: Virtual Machine)を生成し、各VMにおいて個々のOSを動作させることができる。

【0003】

また、プロセッサ資源分割管理機構(PRMF: Processor Resource Management Feature)などのハードウェア機構の補助により、仮想計算機の状態を高速に退避回復する機能を有する仮想計算機システムもある。いずれも、物理的に1台の計算機システムの資源を論理的に分割し、複数のOSを動作させることを可能にする計算機システムである。

【0004】

仮想計算機システム利用の目的としては、異なる種類のOS、異なる設定のOS、異なるバージョンのOSを動作させ、1台の計算機システムとしての運用性を向上させることにある。

【0005】

計算機システムの運用性を向上する方法としては、以下に示すような第1の統

合システム自動運転方法がある。この第1の統合システム自動運転方法では、仮想計算機システムが論理的に複数の仮想計算機に分割され、各仮想計算機において、現用系OSと待機系OSを一台の計算機システムで動作させて、システム切り替えを自動化する。即ち、現用系VMと待機系VMとを用意している。

【0006】

この第1の従来技術における自動運転方法は、ホットスタンバイシステムにおけるOSが使用する主記憶容量について考慮されておらず、待機系OSと現用系OSの双方に固定的に同容量の主記憶を割当てたため、その結果アプリケーションプログラムが動作するための資源を待機系OS用に割当てているため常時使用しない資源が無駄になる問題がある。

【0007】

また、第1の従来技術における自動運転方法は、単一の仮想計算機システムであってもホットスタンバイシステムを構築するためには、各OSの動作を監視するシステム切替え装置を二重系のシステムの各OSに接続する必要があった。

【0008】

仮想計算機への主記憶装置を動的に割当てする方法は、特開平6-110715号公報記載のようなものがある。即ち、各OSの運用時間帯や各システム障害に応じて、イベントの発生により、予め決められた主記憶領域を各VMに割り付け直すことができる。

【0009】

この第2の従来技術における動的に主記憶を再割当てする方法は、同一の仮想計算機システム上で現用系OSと待機系OSを実行することについて考慮されておらず、一方のOS資源を他方のOSに切り替えるためには、双方のOSが正常に動作しなければ、資源を移動できない問題もあった。

【0010】

また、第2の従来技術における動的に主記憶を再割当てする方法は、各仮想計算機上で動作するアプリケーションプログラムを好適に再実行するための主記憶容量について考慮されておらず、障害が発生したアプリケーションプログラムの資源に係わらずVMに割当てた主記憶容量を切り替えるため（障害が発生したアプ

リケーションプログラムに使用される資源以上に資源が切り替わる可能性があるため)、同一OS上で動作する他のアプリケーションプログラムの実行を妨げる問題があった。

【0011】

【発明が解決しようとする課題】

本発明の課題は、同一の仮想記憶計算機システムを利用して、ホットスタンバイシステムを構築したとき、このときのホットスタンバイに要する資源を削減することである。

【0012】

更には、同一の仮想記憶計算機システムを利用して、ホットスタンバイシステムを構築し、現用系OS上の重要アプリケーションプログラムが異常停止した場合の再実行のために用意する資源を削減することである。

【0013】

したがって、本発明の目的は、同一の仮想計算機システムを利用したホットスタンバイシステムにおいて、現用系OS(現用系VM)および待機系OS(待機系VM)に資源を効率的に割当てて仮想計算機システムを提供することにある。

【0014】

また、本発明の別の目的は、同一の仮想計算機システムを利用したホットスタンバイシステムにおいて、障害発生時に、現用系OS上で動作するアプリケーションプログラムを効率的に待機系OS上に切り替える仮想計算機システムを提供することにある。

【0015】

【課題を解決するための手段】

上記課題を解決するためには、待機系OSに割当てて資源を減らし、現用系OSに割当てて資源を増やすことにより、現用系OS上で動作するアプリケーションプログラムにより多くの資源を割当てて、現用系のOSまたはあるアプリケーションプログラムに障害が発生すると現用系OSに割り当てられた資源の一部または全部を待機系OSに割り当てて、

【 0 0 1 6 】

【発明の実施の形態】

本発明の実施例において、現用系OS（現用系VM上で動作するOS）が正常に動作している場合には待機系OSよりも多くの資源を割り当てる。そして、現用系OSでは、アプリケーションプログラムを起動するたびに、必要となる資源を計算し資源を割り当てると共に、自現用系OSが保有する資源が不足する場合には、仮想計算機モニタに連絡し資源拡張をする。

【 0 0 1 7 】

また、現用系OSが異常停止したことを検知すると、現用系OSが使用した資源を回収し、その資源を待機系OS（待機系VM上で動作するOS）に割り当てる。

【 0 0 1 8 】

更に、現用系OS（現用系VM）に割り当てた主記憶の利用用途を記録し、現用系OS上で動作するアプリケーションプログラムが異常停止した場合には、記録した利用用途別に、待機系OS（待機系VM）に割り当て直すことにより、待機系OS上でアプリケーションプログラムを再実行する。

【 0 0 1 9 】

以下、本発明の仮想計算機の資源割当て方法に関する一実施例を、図を用いて詳細に説明する。

図2は、本発明における仮想計算機システム（VMS）のハードウェア構成を示している。

【 0 0 2 0 】

仮想計算機システム50は、ひとつ以上の中央処理装置10-1、10-2、主記憶装置20および入出力制御装置40を含む計算機システムである。仮想計算機システム50には、通信制御装置60および補助記憶装置70が接続されている。

【 0 0 2 1 】

中央処理装置10は、主記憶装置20に格納された命令語を解釈実行するプロセッサである。主記憶装置20は、中央処理装置10-1、10-2にプログラムやデータを供給するための記憶装置である。入出力制御装置40は、補助記憶

装置 7 0 や通信制御装置 6 0 と主記憶装置 2 0 間でプログラムやデータを転送するための装置である。補助記憶装置 7 0 は、プログラムやデータを格納する不揮発の記録装置である。通信制御装置 6 0 は、通信により他の計算機システムにデータを転送するための制御装置である。仮想計算機システム 5 0 では、仮想計算機を管理するソフトウェアまたはファームウェアによって、主記憶装置 2 0 を現用系オペレーティングシステム用領域（現用系 OS 領域）3 0 - 1、待機系オペレーティングシステム用領域（待機系 OS 領域）3 0 - 2、仮想計算機システム 5 0 を管理するための仮想計算機モニタ領域 3 0 - 3 および未割当て領域 3 0 - 4 に分割管理している。

【 0 0 2 2 】

図 1 は、本発明における仮想計算機システム（VMS）のソフトウェア構成を示している。

【 0 0 2 3 】

仮想計算機システム 5 0 上では、計算機システムの資源を管理する仮想計算機モニタ 2 0 0、仮想計算機モニタ 2 0 0 により割当てられた資源を元にアプリケーションプログラム A 4 0 0 - 1 およびアプリケーションプログラム B 4 1 0 - 1 の実行を制御する現用系オペレーティングシステム（現用系 OS）1 0 0 - 1 および現用系 OS が異常を起こした場合に補助記憶装置 7 0 に格納されたアプリケーションプログラム A 4 0 0 - 2 またはアプリケーションプログラム B 4 1 0 - 2 に資源を割当て実行する待機系オペレーティングシステム（待機系 OS）1 0 0 - 2 が動作している。現用系 OS 1 0 0 - 1 および待機系 OS 1 0 0 - 2 は、同系の OS であっても無くても良い。現用系 OS 1 0 0 - 1 と待機系 OS 1 0 0 - 2 が同系の OS である場合には、アプリケーションプログラム A 4 0 0 - 1 とアプリケーションプログラム A 4 0 0 - 2 のプログラムは同じであっても良い（少なくとも同様の業務を遂行するプログラムである）。同様にアプリケーションプログラム B 4 1 0 - 1 とアプリケーションプログラム B 4 1 0 - 2 のプログラムは同じであっても良い（少なくとも同様の業務を遂行するプログラムである）。つまり、アプリケーションプログラム A 4 0 0 - 2 は、アプリケーションプログラム A 4 0 0 - 1 のホットスタンバイ用プログラムであり、アプリケーションプ

ログラムB4 1 0 - 2 は、アプリケーションプログラムB4 1 0 - 1 のホットスタンバイ用プログラムである。

【0 0 2 4】

仮想計算機モニタ2 0 0 は、仮想計算機システム5 0 の資源を管理している。具体的には、仮想計算機システム上で動作する現用系OS 1 0 0 - 1、待機系OS 1 0 0 - 2 の管理、アプリケーションプログラムA 4 0 0 - 1 アプリケーションプログラムB 4 0 0 - 1 を実行する中央処理装置1 0 - 1、1 0 - 2 の割当て、主記憶装置2 0 の特定の割当て単位での管理、及び現用系OS 用領域3 0 - 1 待機系OS 3 0 - 2 仮想計算機モニタ用領域3 0 - 3 および未使用領域3 0 - 4 の管理をしている。

【0 0 2 5】

仮想計算機モニタ2 0 0 の中央処理装置1 0 - 1、1 0 - 2 の割当てについては、特開平9 - 2 6 8 8 9 に記載の様な方法がある。即ち、OS が外部条件の変化に応じて特定のVMを指示して該プロセッサ割り当て量を変更する指令を発行する手段を持ち、仮想計算機制御プログラムが指定されたVMのプロセッサ割当量を変更するものである。

【0 0 2 6】

本発明では、仮想計算機モニタ2 0 0 には、仮想計算機5 0 上で動作するオペレーティングシステムである現用系OS 1 0 0 - 1、待機系OS 1 0 0 - 2 の異常を検知するOS 障害検知処理部2 1 0 を有している。OS 障害検知処理部2 1 0 の検知方法としては、現用系OS 1 0 0 - 1、待機系OS 1 0 0 - 2 を実行中に中央処理装置1 0 - 1、1 0 - 2 の状態が特定の状態(例えば処理が膠着状態になっている)に陥ったことを検知することにより実現できる。

【0 0 2 7】

また、現用系OS 1 0 0 - 1 には、アプリケーションプログラムA 4 0 0 - 1 を実行するために必要な資源を計算し、現用OS 1 0 0 - 1 が持っている資源が不足する場合に仮想計算機モニタ2 0 0 に資源拡張を要求する資源割当て要求処理部1 1 0 - 1、アプリケーションプログラムA 4 0 0 - 1 およびアプリケーションプログラムB 4 1 0 - 1 が障害を発生した場合にその障害度を仮想計算機モ

ニタ 2 0 0 に通知する障害度通知処理部 1 2 0 - 1、仮想計算機モニタ 2 0 0 からの要求により現用系 OS 1 0 0 - 1 が保有する資源を解放する資源切り離し処理部 1 3 0 - 1 および仮想計算機モニタ 2 0 0 の要求により現用系 OS 4 0 0 - 1 に割当てられた資源を有効にする資源追加処理部 1 4 0 - 1 を有している。同様に待機系 OS 1 0 0 - 2 にも資源割当て要求処理部 1 1 0 - 2、障害度通知処理部 1 2 0 - 2、資源切り離し処理部 1 3 0 - 2 および資源追加処理部 1 4 0 - 2 を有している。

【 0 0 2 8 】

以下では、本発明における機能について説明する。

最初に、現用系 OS 1 0 0 - 1 の資源拡張を要求する資源割当て要求処理部 1 1 0 - 1 により、アプリケーションプログラム A 4 0 0 - 1 に資源を割当てる方法を説明する。その具体的な例として、アプリケーションプログラム A 4 0 0 - 1 に、主記憶装置 2 0 の領域からアプリケーションプログラム A 使用域 8 0 - 1 を割当てる方法を示す。また、本説明では、主記憶装置 2 0 を例に取って、説明するが、補助記憶装置 7 0 も同様に資源として扱える。

【 0 0 2 9 】

図 3 は、現用系 OS 1 0 0 - 1 上でアプリケーションプログラム A 4 0 0 - 1 が動作している場合のシステム構成示している。アプリケーションプログラム A 使用域 8 0 - 1 は、主記憶装置 2 0 上の現用系 OS 1 0 0 - 1 が管理する割当てた主記憶領域の一部であり、アプリケーションプログラム A 4 0 0 - 1 のプログラム、データおよび動的な実行情報が格納されている。また、資源管理テーブル 9 0 は、主記憶装置 2 0 の利用目的を管理するテーブルであり、仮想計算機モニタが管理する領域 3 0 - 3 上にある。

【 0 0 3 0 】

アプリケーションプログラム A 使用域 8 0 - 1 の割当てにおいては、現用系 OS 1 0 0 - 1 は、アプリケーションプログラム A 4 0 0 - 1 の実行に際して、ジョブ制御パラメータやシステムパラメータ、ユーザの環境変数等により、アプリケーションプログラム A 4 0 0 - 1 が現用系 OS 1 0 0 - 1 上で動作すると共に障害時に待機系 OS 1 0 0 - 2 上のアプリケーションプログラム A 4 0 0 - 2 に

切り替えられるジョブ(以下ホットスタンバイジョブと略す)であることを検知し、アプリケーションプログラムが実行するために使用する資源量を計算し、現用系OS100-1が保有する資源と比較し、資源が充足しているかを確認する。資源が充足している場合には、現用系OS100-1は、アプリケーションプログラムA400-1が何れの資源を使用しているかを資源割当て要求処理部110-1を使用して仮想計算機モニタ200に通知する。また、資源が不足している場合には、現用系OS100-1は、資源割当て要求処理部110-1を介して、仮想計算機モニタ200から新たな資源を確保する。この資源割当て要求処理部110-1により、仮想計算機モニタ200は、アプリケーションプログラムA400の使用域80-1とそれを管理する情報を資源管理テーブル90に記録する。仮想計算機モニタ200が現用系OS100-1に資源を割当てる方法は、特開平6-110715号公報記載の方法を用いて実現しても良い。特開平6-110715号公報には、拡張しようとするVMのゲスト領域の高位アドレスの領域を未接続状態として、その後その高位アドレスの領域をそのVMの接続領域にして領域の拡大を行なうことが記載されている。また、現用系OS100-1がアプリケーションプログラムA400-1に資源を割当てる方法は、アプリケーションプログラムA使用域80-1から資源を分割して割当てることにより実現できる。

【0031】

従来の仮想計算機上で動作する現用系OS100-1は、仮想計算機モニタ200によって割当てられた資源を分割管理し、アプリケーションプログラムA400-1の実行において、その資源の一部を割当て実行する。具体的には、補助記憶装置70上に格納されたアプリケーションプログラムA400-2は、OS100-1に割当てられた領域30-1上に展開されて、中央処理装置10-1, 10-2によって実行される。

【0032】

本発明における現用系OS100-1は、アプリケーションプログラムA400-1の実行に先立ち、現用系OS100-1が保有する資源が十分か否かを判定し、不足する場合には、仮想計算機モニタ200に対して、資源の拡張を要求

する。仮想計算機モニタ200は、仮想計算機間の資源の割当てを調停し、その結果、現用系OS100-1に資源の拡張が可能であれば、仮想計算モニタ200が所有する資源の一部を現用系OS100-1に追加する。これにより、現用系OS100-1上でアプリケーションプログラムA400-1が好適な資源を確保して実行が可能になる。

【0033】

図4は、資源管理テーブル90の構成とアプリケーションプログラム実行におけるテーブルの状態変化を示している。個々の状態は、状態Aがアプリケーションプログラムの未実行状態、状態BがホットスタンバイのジョブであるアプリケーションプログラムAを実行した場合、CがBの状態に対し非ホットスタンバイジョブであるアプリケーションプログラムBを実行した場合の状態を示している。資源管理テーブル90は、個々の領域を識別する領域名称、使用量、使用者および異常時に資源を切り替える先を記録したエントリ91-1、91-2、91-3、91-nから構成されている。

【0034】

アプリケーションプログラムは総てOSに障害が発生したとき待機系によって実行が引き継がれるものではない。予め定められた重要なアプリケーションプログラムのみが待機系のOSの制御の下に実行が継承される。OSに障害が生じた場合は予め決められたアプリケーションプログラム以外の処理は中断され、その使用記憶領域は仮想計算機モニタに戻される。ある重要アプリケーションプログラムにのみに障害が生じれば、そのアプリケーションプログラムは待機系OSの制御の下で実行が継続され、他のアプリケーションプログラムは現用系OSの制御の下で実行がそのまま続けられる。

【0035】

状態Aから状態Bへの変化は、現用系OS100-1には、当初160MB割当てられており、128MBの資源を使用するホットスタンバイ用のアプリケーションプログラムA400-1を実行した場合に、資源管理テーブルには、現用系OS使用量を128MB減算し、ホットスタンバイのために切り替え先のOSが定義される。

【 0 0 3 6 】

状態Bから状態Cへの変化では、現用系OS 1 0 0 - 1には、アプリケーションプログラムBを実行するために十分な資源を保有していなかった。したがって、現用系OS 1 0 0 - 1が仮想計算機モニタ2 0 0に資源割り当て要求を出し、仮想計算機システムの未使用領域を確保して、現用系OS 1 0 0 - 1上で動作させている。異常時切り替え先は、現用系OS 1 0 0 - 1またはアプリケーションプログラムA 4 0 0 - 1が異常停止時にその資源をどこの領域に再割り当てするかを示している。

【 0 0 3 7 】

次に、図5を用いて、現用系OS 1 0 0 - 1が障害になった場合の処理フローを例に資源追加処理部1 4 0 - 2を説明する。仮想計算機モニタ2 0 0は、OS障害検知処理部2 1 0により現用系OS 1 0 0 - 1の障害を検知し（ステップ5 0 1）、領域管理テーブル9 0の各エントリ9 1 - 1, 9 1 - 2, . . . 9 1 - nを参照することにより現用OS 1 0 0 - 1が保有する資源の移行先を決定し（ステップ5 0 2）、領域管理テーブル9 0の異常時切り替え先に従って待機系OS 1 0 0 - 2に通知する（ステップ5 0 3）。通知を受け取った待機系OS 1 0 0 - 2は追加された資源を資源追加処理部1 4 0 - 2を呼び出すことによりOS資源に組み入れ（ステップ5 0 4）、仮想計算機モニタ2 0 0に資源の組み入れが完了したことを通知する（ステップ5 0 5）。その通知を受け取った仮想計算機モニタ2 0 0は領域管理テーブル9 0の各エントリ9 1 - 1, 9 1 - 2, . . . 9 1 - nの使用者を更新し（ステップ5 0 5）、各アプリケーションプログラムが動作していた領域については待機系OS 1 0 0 - 2にアプリケーションプログラムA 4 0 0 - 2の起動を通知する（ステップ5 0 6）。これにより、待機系OS 1 0 0 - 2は十分な資源を持って、アプリケーションプログラムA 4 0 0 - 2を実行できる。

【 0 0 3 8 】

本実施例では、待機系OS 1 0 0 - 2の資源追加処理部1 4 0 - 2は、オペレーティングシステムが使用できる主記憶量が増大したことを示しており、それまで使用不可にしていた主記憶アドレスを使用可能にすることにより実現できる。

【 0 0 3 9 】

次に、図 6 を用いて、現用系 OS 1 0 0 - 1 上で動作するアプリケーションプログラム A 4 0 0 - 1 のみが障害になった場合の処理フローを例に障害度通知処理部 1 2 0 - 1 および資源切り離し処理部 1 3 0 - 1 を説明する。現用系 OS 1 0 0 - 1 がアプリケーションプログラム 4 0 0 - 1 の障害を検知し（ステップ 6 0 1）、アプリケーションプログラムの障害度を仮想計算機モニタ 2 0 0 に通知する（ステップ 6 0 2）。

【 0 0 4 0 】

仮想計算機モニタ 2 0 0 では、障害度を判別し（ステップ 6 0 3）、領域管理テーブル 9 0 の各エントリ 9 1 - 1, 9 1 - 2, . . . 9 1 - n を参照することによりアプリケーションプログラム 4 0 0 - 1 が保有する資源の移行先を決定し（ステップ 6 0 4）、現用系 OS 1 0 0 - 1 にアプリケーションプログラム A 4 0 0 - 1 が使用している資源の切り離しを通知する（ステップ 6 0 5）。

【 0 0 4 1 】

現用系 OS 1 0 0 - 1 では資源切り離し処理部 1 3 0 - 1 を呼び出して資源を切り離し（ステップ 6 0 6）、資源の切り離し完了を仮想計算機モニタ 2 0 0 に通知する（ステップ 6 0 7）。

【 0 0 4 2 】

仮想計算機モニタ 2 0 0 では、領域管理テーブル 9 0 の各エントリ 9 1 - 1, 9 1 - 2, . . . 9 1 - n を参照することによりアプリケーションプログラム 4 0 0 - 1 の異常時切り替え先に従って待機系 OS 1 0 0 - 2 に通知する（ステップ 6 0 8）。

【 0 0 4 3 】

通知を受け取った待機系 OS 1 0 0 - 2 は追加された資源を資源追加処理部 1 4 0 - 2 を呼び出すことにより OS 資源に組み入れ（ステップ 6 0 9）、仮想計算機モニタ 2 0 0 に資源の組み入れが完了したことを通知する（ステップ 6 1 0）。

【 0 0 4 4 】

その通知を受け取った仮想計算機モニタ 2 0 0 は領域管理テーブル 9 0 の各エ

ントリ 9 1 - 1, 9 1 - 2, . . . 9 1 - n の使用者を更新し (ステップ 6 1 1)、アプリケーションプログラムが動作していた領域については待機系 OS 1 0 0 - 2 にアプリケーションプログラム A 4 0 0 - 2 の起動を通知する (ステップ 6 1 2)。

【0 0 4 5】

仮想計算機モニタ 2 0 0 におけるアプリケーションプログラム 4 0 0 - 1 の障害度の判定においては、現用系 OS 1 0 0 - 1 が独自に回復できる軽度の障害については資源を移動せずにアプリケーションプログラムを再起動させる。また、現用系 OS 1 0 0 - 1 が独自に回復できる軽度の障害であっても現用系 OS 1 0 0 - 1 による障害回復が難しい場合には、待機系 OS 1 0 0 - 2 に切り替えることにより短時間で再起動が可能になる。

【0 0 4 6】

以上のようにして、本実施例によれば、複数のオペレーティングシステム間で動的に資源変更が可能な仮想計算機システムを提供できる。また、アプリケーションプログラム単位に動的に資源変更が可能な仮想計算機システムを提供できる。

【0 0 4 7】

さらに、現用系 OS 1 0 0 - 1 が使用する領域がアクセス不可になるような重度障害を発生した場合、仮想計算機モニタ 2 0 0 に小容量の未使用領域を確保しておくことにより、使用不可になった資源の一部を切り離して、未使用領域の一部を加えることにより、待機系 OS においてアプリケーションプログラムを再起動することが可能な仮想計算機システムを提供できる。

【0 0 4 8】

さらに、計算機システムが動的に資源を追加可能になる場合に、仮想計算機モニタに資源を一時的に保持することにより、ホットスタンバイ運用を停止させることなく計算機の資源を追加することが可能な仮想計算機システムを提供できる。

【 0 0 4 9 】

【発明の効果】

本発明によれば、オペレーティングシステム間で仮想計算機システムの資源が流動化することが可能になり、仮想計算機システムに必要な資源が削減できる。

【図面の簡単な説明】

【図 1】 本発明によるソフトウェア構成を示す。

【図 2】 本発明によるハードウェア構成を示す。

【図 3】 現用系オペレーティングシステム上でホットスタンバイのアプリケーションを動作させた場合のシステム構成を示す。

【図 4】 資源管理テーブルの説明を示す。

【図 5】 現用系 OS 障害時の資源移行処理フローを示す。

【図 6】 ホットスタンバイのアプリケーションプログラムが障害時の資源移行フローを示す

【符号の説明】

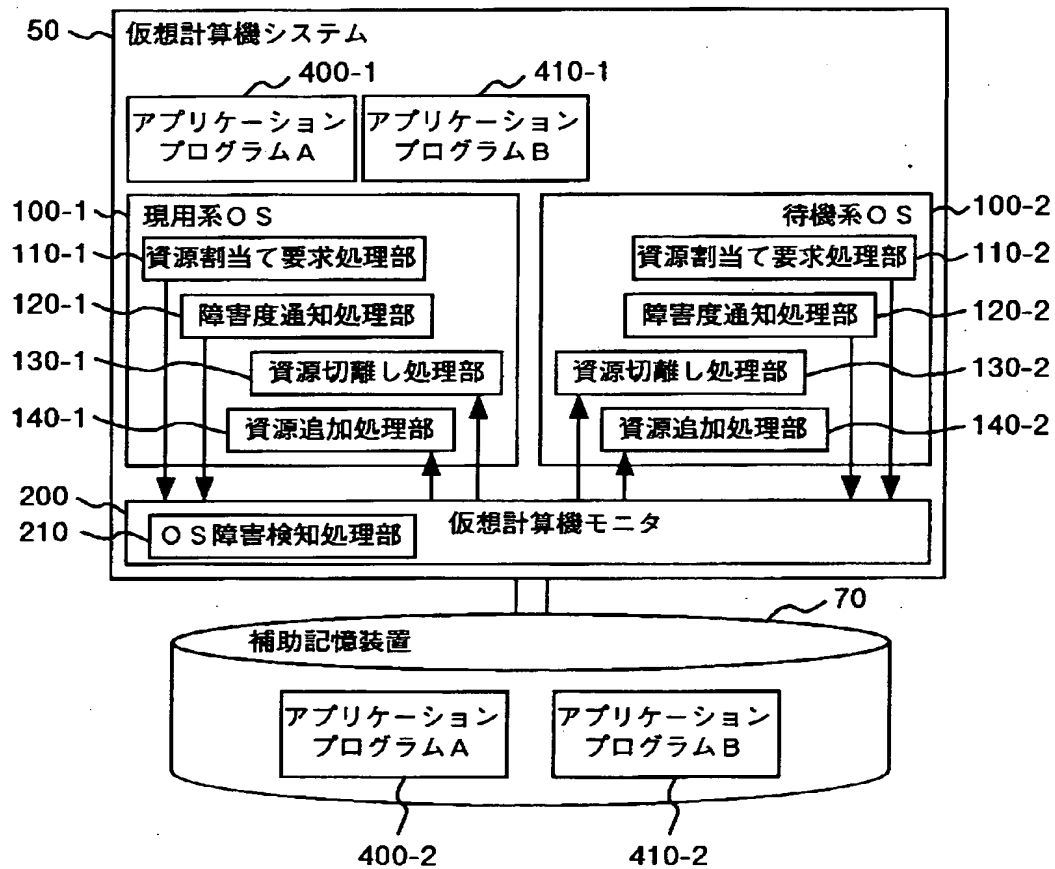
1 0 - 1, 1 0 - 2 … 中央処理装置、 2 0 … 主記憶装置、 4 0 … 入出力制御装置、 5 0 … 仮想計算機システム、 6 0 … 通信制御装置、 7 0 … 補助記憶装置、 3 0 - 1, 3 0 - 2, 3 0 - 3, 8 0 - 1 … 各種主記憶使用領域、 1 0 0 - 1, 1 0 0 - 2 … オペレーティングシステム、 2 0 0 … 仮想計算機モニタ、 4 0 0 - 1, 4 0 0 - 2, 4 1 0 - 1, 4 1 0 - 2 … アプリケーションプログラム

【書類名】 図面

【図 1】

図 1

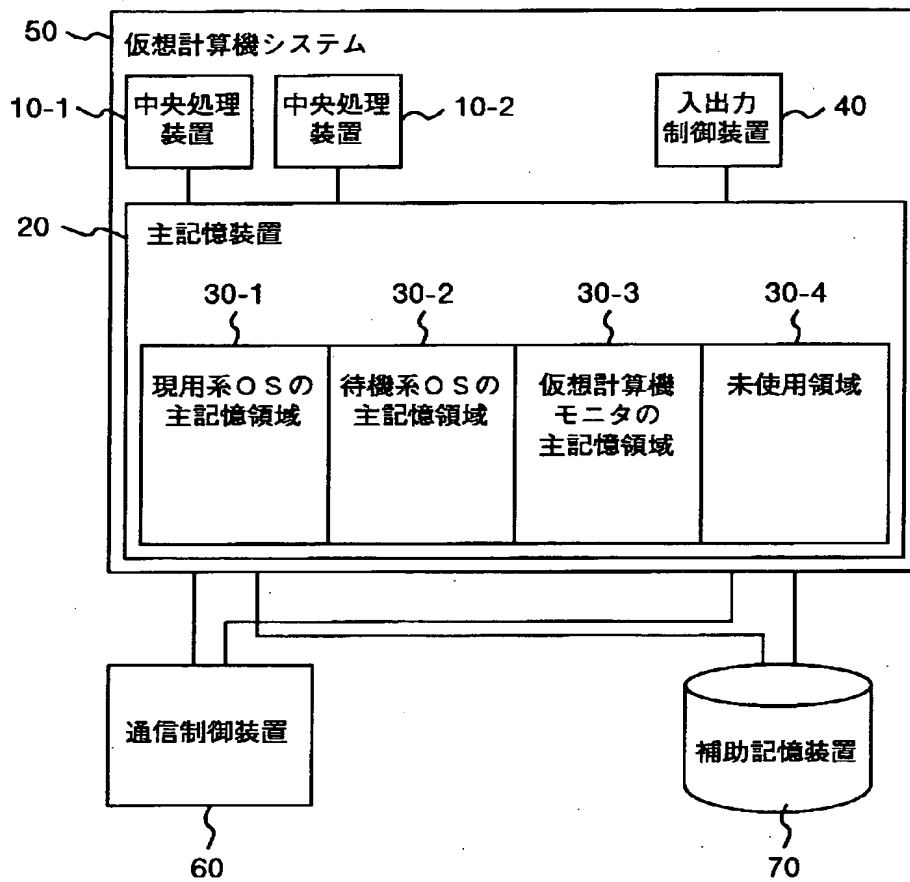
ソフトウェア構成



【図 2】

図 2

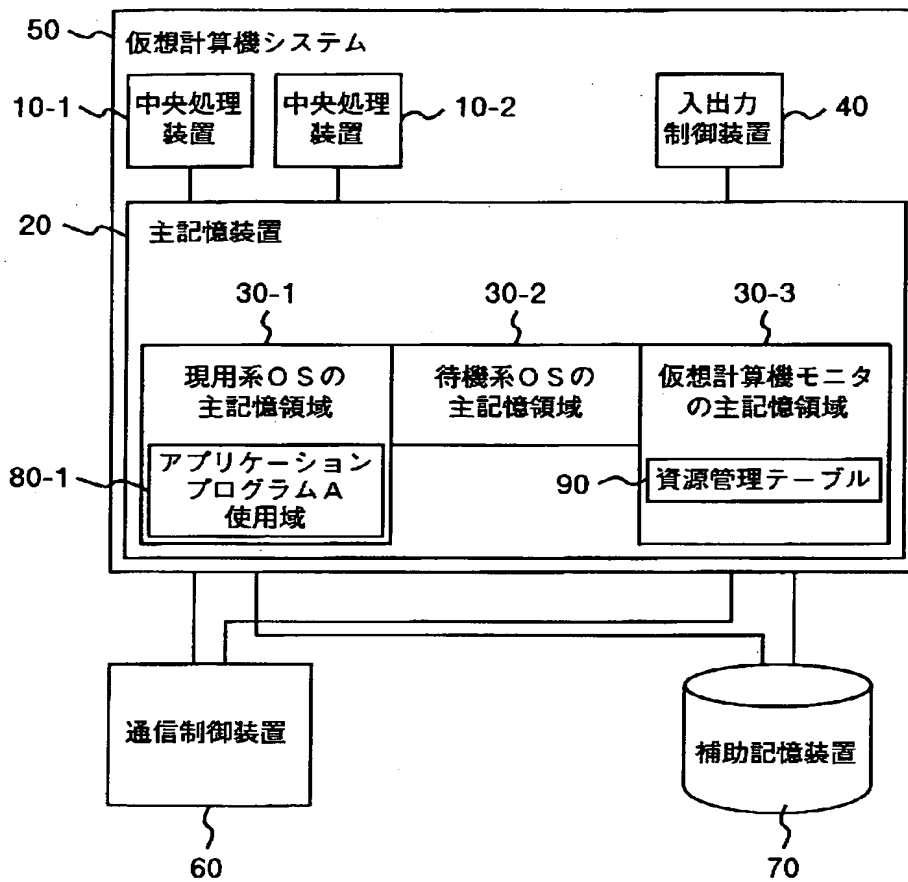
ハードウェア構成



【図 3】

図 3

現用系動作時のシステム構成



【図 4】

図 4

資源管理テーブルの説明図

状態 A				
	領域名称	使用量	使用者	異常時切り替え先
91-1 ~	現用系 OS 使用域	160MB	現用 OS	現用 OS
91-2 ~	待機系 OS 使用域	32MB	待機 OS	待機 OS
91-3 ~	仮想モニタ使用域	32MB	モニタ	モニタ
91-n ~	未使用域	32MB	モニタ	モニタ

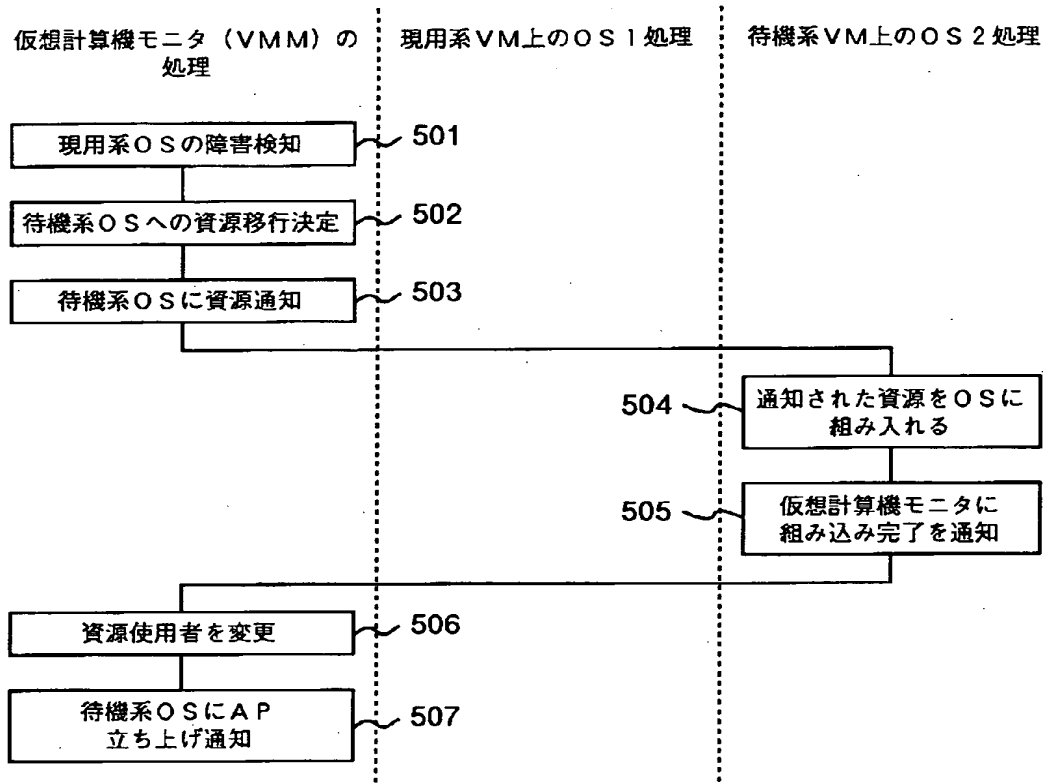
状態 B				
	領域名称	使用量	使用者	異常時切り替え先
91-1 ~	現用系 OS 使用域	32MB	現用 OS	現用 OS
91-2 ~	待機系 OS 使用域	32MB	待機 OS	待機 OS
91-3 ~	仮想モニタ使用域	32MB	モニタ	モニタ
91-4 ~	アプリケーションプログラム A 使用域	128MB	現用 OS	待機 OS
91-n ~	未使用域	32MB	モニタ	モニタ

状態 C				
	領域名称	使用量	使用者	異常時切り替え先
91-1 ~	現用系 OS 使用域	32MB	現用 OS	現用 OS
91-2 ~	待機系 OS 使用域	32MB	待機 OS	待機 OS
91-3 ~	仮想モニタ使用域	32MB	モニタ	モニタ
91-4 ~	アプリケーションプログラム A 使用域	128MB	現用 OS	待機 OS
91-5 ~	アプリケーションプログラム B 使用域	32MB	現用 OS	モニタ
91-n ~	未使用域	0MB	モニタ	モニタ

【図 5】

図 5

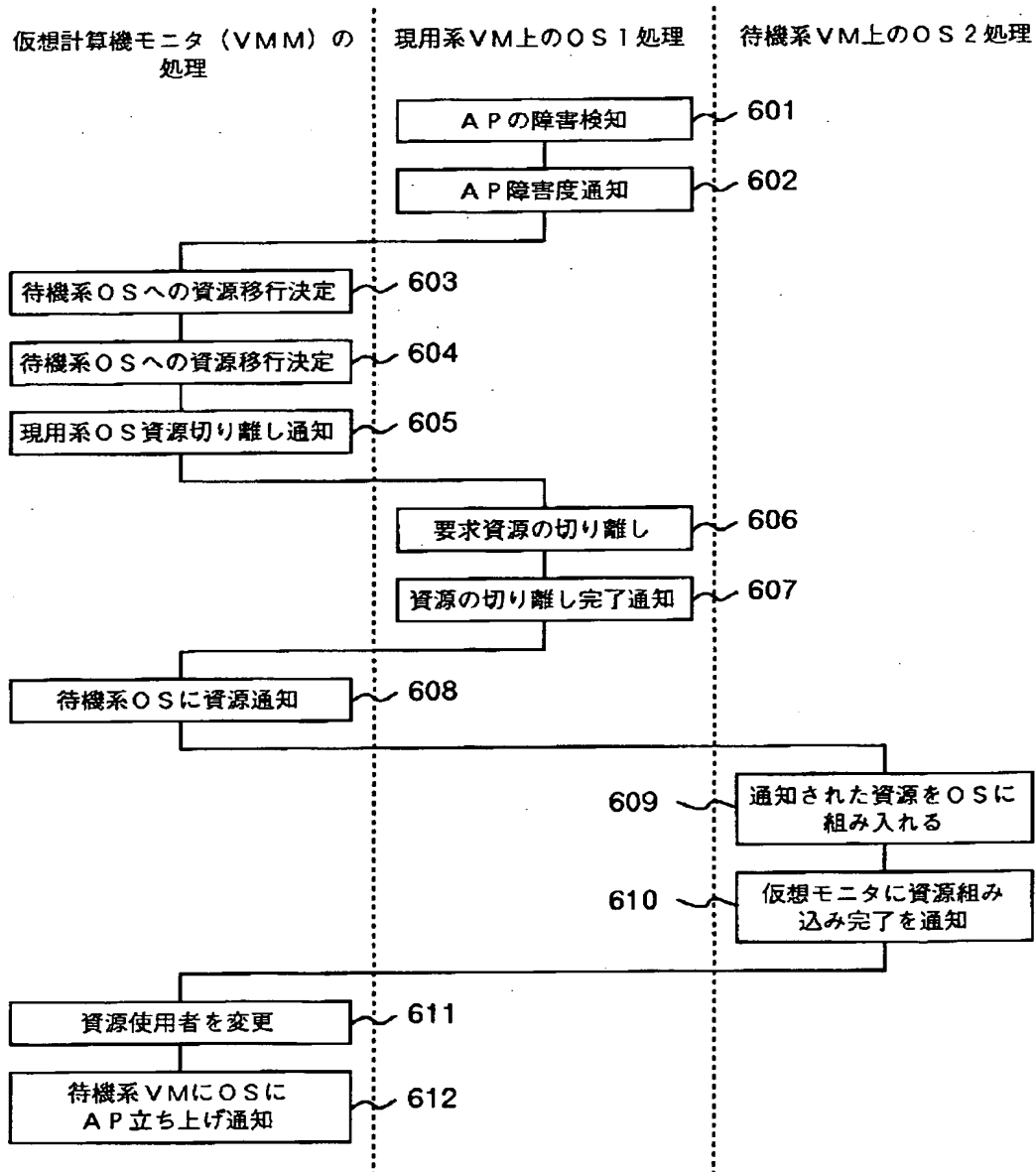
現用系OS障害時の処理フロー



【図 6】

図 6

アプリケーションプログラム障害時のフロー



【書類名】 要約書

【要約】

【課題】

本発明は、単一の計算機システムにおいて、ホットスタンバイ構成を組み複数のオペレーティングシステムおよびアプリケーションプログラムを効率的に制御する仮想計算機システムを提供することにある。

【解決手段】

仮想計算機システムにおいて、複数のオペレーティングシステムを起動した場合、個々のオペレーティングシステムに割当て主記憶量を少なくすることにより、固定的に割当てられる主記憶量を減らし、アプリケーションプログラムを実行する場合に仮想計算機モニタに資源の拡張を要求して確保することにより、オペレーティングシステムが新たな資源を確保し、アプリケーションプログラムを実行する。次に、仮想計算機システムを用いてホットスタンバイシステムを構築する場合、現用系オペレーティングシステムの資源を待機系オペレーティングシステムに再割り当てすることにより、資源の再利用を図る。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所